2016

Strongly convex functions, Moreau envelopes and the generic nature of convex functions with strong minimizers

Chayne Planiden
University of British Columbia, Okanagan, chayne@uow.edu.au

Xianfu Wang
University of British Columbia, Okanagan, shawn.wang@ubc.ca

Strongly convex functions, Moreau envelopes and the generic nature of convex functions with strong minimizers

Abstract

In this work, using Moreau envelopes, we define a complete metric for the set of proper lower semicontinuous convex functions. Under this metric, the convergence of each sequence of convex functions is epiconvergence. We show that the set of strongly convex functions is dense but it is only of the first category. On the other hand, it is shown that the set of convex functions with strong minima is of the second category.

Disciplines

Engineering | Science and Technology Studies

Publication Details

Planiden, C. & Wang, X. (2016). Strongly convex functions, Moreau envelopes and the generic nature of convex functions with strong minimizers. SIAM Journal on Optimization, 26 (2), 1341-1364.

Strongly convex functions, Moreau envelopes and the generic nature of convex functions with strong minimizers

C. Planiden* X. Wang†

April 13, 2016

Abstract

In this work, using Moreau envelopes, we define a complete metric for the set of proper lower semicontinuous convex functions in a finite-dimensional space. Under this metric, the convergence of each sequence of convex functions is epi-convergence. We show that the set of strongly convex functions is dense but it is only of the first category. On the other hand, it is shown that the set of convex functions with strong minima is of the second category.

AMS Subject Classification: Primary 54E52, 52A41, 90C25; Secondary 49K40.

Keywords: Attouch-Wets metric, Baire category, complete metric space, convex function, epi-convergence, epi-topology, generic set, meager set, Moreau envelope, proximal mapping, strong minimizer, strongly convex.

1 Introduction

Minimizing convex functions is fundamental in optimization, both in theory and in algorithm design. For most applications, the assertions that can be made about a class of convex functions are of greater value than those concerning a particular problem. This theoretical analysis is valuable for the insights gained on the behaviour of the entire class of functions. Our main result in this paper states that in $\mathbb{R}^n$, the set of all proper lower semicontinuous (lsc) convex functions that have strong minimizers is of second category. Studying strong minima is important, because numerical methods usually produce asymptotically minimizing sequences, and we can assert convergence of asymptotically minimizing sequences when the function has a strong minimizer. As shown in

*Mathematics, University of British Columbia Okanagan, Kelowna, B.C. V1V 1V7, Canada. Research by this author was supported by UBC UGF and by NSERC of Canada. chayne.planiden@alumni.ubc.ca.

†Mathematics, University of British Columbia Okanagan, Kelowna, B.C. V1V 1V7, Canada. Research by this author was partially supported by an NSERC Discovery Grant. shawn.wang@ubc.ca.

[25], the proximal mapping of a strongly convex function is a contraction, and the proximal point method converges at a linear rate. The strongly convex function is also of great use in optimization problems, as it can significantly increase the rate of convergence of first-order methods such as projected subgradient descent [18], or more generally the forward-backward algorithm [5, Example 27.12]. Although every strongly convex function has a strong minimizer, we show that the set of strongly convex functions is only of the first category.

As a proper lsc convex function allows the value infinity, we propose to relate the function to its Moreau envelope. The importance of the Moreau envelope in optimization is clear; it is a regularizing (smoothing) function [20, 21], and in the convex setting it has the same local minima and minimizers as its objective function [27, 30].

The key tool we use is Baire category. A property is said to be generic if it holds for a second category set. We will work in a metric space defined by Moreau envelopes. In this setting, there are many nice properties of the set of Moreau envelopes of proper, lsc, convex functions. This set is proved to be closed and convex. Moreover, as a mapping from the set of proper lsc convex functions to the set of Moreau envelopes of convex functions, the Moreau envelope mapping is bijective. We provide a detailed analysis of functions with strong minima, strongly convex functions, and their Moreau envelopes.

The organization of the present work is the following. Section 2 contains notation and definitions, as well as some preliminary facts and lemmas about Baire category, epi-convergence of convex functions, strongly convex functions and strong minimizers that we need to prove the main results. We show that the Moreau envelope of a convex function inherits many nice properties of the convex function, such as coercivity and strong convexity. In Section 3, using Moreau envelopes of convex functions, we propose to use the Attouch-Wets metric on the set of proper lsc convex functions. It turns out that this metric space is complete, and it is isometric to the metric space of Moreau envelopes endowed with uniform convergence on bounded sets. The main results of this paper are presented in Section 4. We give some characterizations of strong minimizers of convex functions, that are essential for our Baire category approach. We establish Baire category classification of the sets of strongly convex functions, convex functions with strong minima, and convex coercive functions. Our main result says that most convex functions have strong minima, which in turn implies that the set of convex functions not having strong minimizers is small. Surprisingly, the set of strongly convex functions is only of the first category. In addition, we show that a convex function is strongly convex if and only if its proximal mapping is a down-scaled proximal mapping. Concluding remarks and areas of future research are mentioned in Section 5.

A comparison to literature is in order. In [29], Baire category theory was used to show that most (i.e. a generic set) maximally monotone operators have a unique zero. In [22], a similar track was taken, but it uses the perspective of proximal mappings in particular, ultimately proving that most classes of convex functions have a unique minimizer. The technique of this paper differs in that it is based on functions. We use Moreau envelopes of convex functions, strong minimizers and strongly convex functions rather than subdifferentials. While Beer and Lucchetti obtained a similar result on generic well-posedness of convex optimization, their approach relies on epigraphs of convex functions [7, 8, 9]. Our Moreau envelope approach is more accessible and natural to practical optimizers, because taking the Moreau envelope is a popular regularization method used in the optimization community. We also give a systematic study of strongly convex functions, a special subclass of uniformly convex functions, which is new to the best of our knowledge. In [16], the definition of generic Tikhonov well-posedness of convex problems is given. There, it is assumed that either the convex functions are finite-valued and the set of convex functions is equipped with uniform convergence on bounded sets, or the convex functions are all continuous on the whole space. See also [28] for generic nature of constrained optimization problems, and [7, 16, 19, 24] for well-posedness in optimization. For comprehensive generic results on fixed points of nonexpansive mappings and firmly nonexpansive mappings, we refer the reader to [23].

2 Preliminaries

2.1 Notation

All functions in this paper have their domains in $\mathbb{R}^n$, where $\mathbb{R}^n$ is Euclidean space equipped with inner product $\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i$, and induced norm $\|x\| := \sqrt{\langle x, x\rangle}$. The extended real line $\mathbb{R}\cup\{\infty\}$ is denoted $\overline{\mathbb{R}}$. We denote the set of natural numbers by $\mathbb{N}$. We use $\operatorname{dom} f$ for the domain of f, $\operatorname{int}\operatorname{dom} f$ for the interior of the domain of f, $\operatorname{bdry}\operatorname{dom} f$ for the boundary of the domain of f, and $\operatorname{epi} f$ for the epigraph of f. We use $\Gamma_0(X)$ to represent the set of proper lsc convex functions on the space X, with the terms proper, lsc, and convex as defined in [5, 27]. More precisely, f is proper if $-\infty \notin f(X)$ and $\operatorname{dom} f \neq \varnothing$; f is lsc at x if $x_k \to x$ implies $\liminf_{k\to\infty} f(x_k) \geq f(x)$, when this is true at every $x \in X$ we call f lsc on X; f is convex if

$$(\forall x, y \in \operatorname{dom} f)(\forall 0 \leq \alpha \leq 1) \quad f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y).$$

The symbol G_δ is used to indicate a countable intersection of open sets. The identity mapping or matrix is $\operatorname{Id} : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto x$. We use $\mathbb{B}_r(x)$ for the open ball centered at x of radius r, and $\mathbb{B}_r[x]$ for the closed ball. For a set $C \subseteq \mathbb{R}^n$, its closure is denoted by $\overline{C}$. The closed line segment between $x, y \in \mathbb{R}^n$ is $[x, y] := \{\lambda x + (1-\lambda)y : \ 0 \leq \lambda \leq 1\}$. For a sequence of functions $\{f_\gamma\} \subseteq \Gamma_0(\mathbb{R}^n)$, we use $f_\gamma \xrightarrow{p} f$ to indicate pointwise convergence, $f_\gamma \xrightarrow{e} f$ for epi-convergence, and $f_\gamma \xrightarrow{u} f$ for uniform convergence to the function f.

2.2 Baire category

Let (X, d) be a metric space, where X is a set and d is a metric on X.

Definition 2.1. *A set $S \subseteq X$ is* dense *in X if every element of X is either in S, or a limit point of S. A set is* nowhere dense *in X if the interior of its closure in X is empty.*

Definition 2.2. *A set $S \subseteq X$ is* of first category (meager) *if S is a union of countably many nowhere dense sets. A set $S \subseteq X$ is* of second category (generic) *if $X \setminus S$ is of first category.*

The following Baire category theorem is essential for this paper.

Fact 2.3 (Baire)**.** ([32, Theorem 1.47] or [5, Corollary 1.44]) *Let (X, d) be a complete metric space. Then any countable intersection of dense open subsets of X is dense.*

Fact 2.4. *Finite-dimensional space $\mathbb{R}^n$ is separable. That is, $\mathbb{R}^n$ has a countable subset that is dense in $\mathbb{R}^n$.*

2.3 Convex analysis

In this section we state several key facts about convex functions that we need in order to prove the main results in subsequent sections.

2.3.1 Subdifferentials of convex functions

Let $f \in \Gamma_0(\mathbb{R}^n)$. The set-valued mapping

$$\partial f\colon\ \mathbb{R}^n \rightrightarrows \mathbb{R}^n\colon x \mapsto \left\{x^* \in \mathbb{R}^n \mid (\forall y \in \mathbb{R}^n)\ \langle y - x, x^*\rangle + f(x) \leq f(y)\right\}$$

is the subdifferential operator of f.

Fact 2.5. [5, Theorem 20.40] *If $f \in \Gamma_0(\mathbb{R}^n)$, then ∂f is maximally monotone.*

Fact 2.6. ([27, Theorem 12.41], [3, Theorem 2.51]) *For any maximally monotone mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, the set $\operatorname{dom} T$ is almost convex. That is, there exists a convex set $C \subseteq \mathbb{R}^n$ such that $C \subseteq \operatorname{dom} T \subseteq \overline{C}$. The same applies to the set $\operatorname{ran} T$.*

Definition 2.7. *The* Fenchel conjugate *of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined as*

$$(\forall v \in \mathbb{R}^n) \qquad f^*(v) := \sup_{x \in \mathbb{R}^n} \{\langle v, x\rangle - f(x)\}.$$

Fact 2.8. [26, Corollary 23.5.1] *If $f \in \Gamma_0(\mathbb{R}^n)$, then ∂f^* is the inverse of ∂f in the sense of multivalued mappings, i.e. $x \in \partial f^*(x^*)$ if and only if $x^* \in \partial f(x)$.*

2.3.2 Convex functions and their Moreau envelopes

Definition 2.9. *The* Moreau envelope *of a proper, lsc function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined as*

$$(\forall x \in \mathbb{R}^n) \qquad e_\lambda f(x) := \inf_y \left\{ f(y) + \frac{1}{2\lambda}\|y - x\|^2 \right\}.$$

The associated proximal mapping *is the (possibly empty) set of points at which this infimum is achieved, and is denoted* $\operatorname{Prox}_f^\lambda$:

$$(\forall x \in \mathbb{R}^n) \qquad \operatorname{Prox}_f^\lambda(x) := \operatorname*{argmin}_y \left\{ f(y) + \frac{1}{2\lambda}\|y - x\|^2 \right\}.$$

In terms of the subdifferential of f, the proximal mapping $\mathrm{Prox}_f^1 = (\mathrm{Id} + \partial f)^{-1}$. In this paper, without loss of generality we use $\lambda = 1$. The theory developed here is equally applicable with any other choice of $\lambda > 0$.

Fact 2.10. ([5, Proposition 12.29] or [27, Theorem 2.26]) *Let $f \in \Gamma_0(\mathbb{R}^n)$. Then $e_1 f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable on $\mathbb{R}^n$, and its gradient*

$$\nabla e_1 f = \mathrm{Id} - \mathrm{Prox}_f^1$$

is 1-Lipschitz continuous, i.e., nonexpansive.

One important concept for studying the convergence of extended-valued functions is that of epi-convergence, see, e.g., [1, 7, 27]. We first remind the reader of the concepts of limit inferior and limit superior for a sequence of sets.

Definition 2.11. *The distance from a point x to a set $S \subset \mathbb{R}^n$ is defined by the* distance function d_S :

$$d_S(x) := \inf_{y \in S} \|x - y\|.$$

Definition 2.12. *The* limit inferior *of a sequence of sets $\{S_\gamma\}_{\gamma=1}^\infty$ is denoted $\liminf_{\gamma \to \infty} S_\gamma$, and is defined*

$$\liminf_{\gamma \to \infty} S_\gamma := \{x \in \mathbb{R}^n : \limsup_{\gamma \to \infty} d_{S_\gamma}(x) = 0\}.$$

Similarly, the limit superior *is denoted $\limsup_{\gamma \to \infty} S_\gamma$, and is defined*

$$\limsup_{\gamma \to \infty} S_\gamma := \{x \in \mathbb{R}^n : \liminf_{\gamma \to \infty} d_{S_\gamma}(x) = 0\}.$$

Definition 2.13. *The* lower epi-limit *of a sequence $\{f_\gamma\}_{\gamma=1}^\infty \subseteq \mathbb{R}^n$ is the function having as its epigraph the outer limit of the sequence of sets* $\operatorname{epi} f_\gamma$:

$$\operatorname{epi}(\operatorname*{eliminf}_{\gamma \to \infty} f_\gamma) := \limsup_{\gamma \to \infty} (\operatorname{epi} f_\gamma).$$

Similarly, the upper epi-limit *of $\{f_\gamma\}_{\gamma=1}^\infty$ is the function having as its epigraph the inner limit of the sets* $\operatorname{epi} f_\gamma$:

$$\operatorname{epi}(\operatorname*{elimsup}_{\gamma \to \infty} f_\gamma) := \liminf_{\gamma \to \infty} (\operatorname{epi} f_\gamma).$$

When these two functions coincide, the epi-limit *is said to exist and the functions are said to* epi-converge *to f. We denote this by using the following notation:*

$$f_\gamma \xrightarrow{e} f \;\; \textit{if and only if} \quad \operatorname{epi} f_\gamma \to \operatorname{epi} f.$$

Definition 2.14. *Let $\{f_\gamma\}$ be a sequence of functions with the same domain D. The sequence converges pointwise to a function f, denoted $f_\gamma \xrightarrow{p} f$, if and only if*

$$(\forall x \in D) \qquad \lim_{\gamma \to \infty} f_\gamma(x) = f(x).$$

We refer the reader to [7, 9, 11, 27] for further details on epi-convergence, e.g., continuity, stability and applications in optimization. The analysis of the limit properties of sequences of convex functions via their Moreau envelopes is highlighted by the following fact.

Fact 2.15. ([27, Theorem 7.37], [1]) *Let* $\{f_\gamma\}_{\gamma=1}^{\infty} \subseteq \Gamma_0(\mathbb{R}^n)$, $f \in \Gamma_0(\mathbb{R}^n)$. *Then*

$$f_\gamma \xrightarrow{e} f \quad \text{if and only if} \quad e_1 f_\gamma \xrightarrow{p} e_1 f.$$

Moreover, the pointwise convergence of $e_1 f_\gamma$ *to* $e_1 f$ *is uniform on all bounded subsets of* $\mathbb{R}^n$*, hence yields epi-convergence to* $e_1 f$ *as well.*

Remark 2.16. *(1). Fact 2.15 would not be valid as it stands in infinite-dimensional space. A uniform limit on bounded sets of Moreau envelopes will remain a Moreau envelope, but this can fail to be true when considering pointwise convergence of Moreau envelopes, see, e.g., [1, Remark 2.71].*

(2). When X *is a Banach space, we refer to Beer [7, p. 263]: for* $\{f_\gamma\}_{\gamma=1}^{\infty} \subseteq \Gamma_0(X)$*, one has that* $f_\gamma \xrightarrow{aw} f_0$ *if and only if* $e_1 f_\gamma \to e_1 f_0$ *uniformly on bounded subsets of* X*. When* X *is infinite dimensional, Attouch-Wets convergence is a stronger concept than epi-convergence, see, e.g., [7, p. 235] or [11, Theorem 6.2.14].*

Two more nice properties about Moreau envelopes are the following.

Fact 2.17. [27, Example 1.46] *For any proper, lsc function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $\inf f = \inf e_1 f$.

Lemma 2.18. [26, Theorem 31.5] *Let* $f \in \Gamma_0(\mathbb{R}^n)$. *Then*

$$(\forall x \in \mathbb{R}^n) \quad e_1 f(x) + e_1 f^*(x) = \frac{1}{2}\|x\|^2.$$

For more properties of Moreau envelopes of functions, we refer the reader to [1, 5, 26, 27].

2.4 Strong minimizers, coercive convex functions and strongly convex functions

We now present some basic properties of strong minimizers, strongly convex functions, and coercive functions.

Definition 2.19. *A function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is said to attain a* strong minimum *at* $\bar{x} \in \mathbb{R}^n$ *if*

(i) $f(\bar{x}) \le f(x)$ *for all* $x \in \operatorname{dom} f$, *and*

(ii) $f(x_k) \to f(\bar{x})$ *implies* $x_k \to \bar{x}$.

In existing literature, a function f having a strong minimizer is also known as $(\mathbb{R}^n, f)$ being *well-posed* in the sense of Tikhonov, as defined by Dontchev and Zolezzi in [16, Chapter I] and used in [7, 13, 15, 17]. [16, Theorem 12] shows that f has a strong minimizer $\bar{x}$ if and only if there exists a forcing function c such that

$$f(x) \geq f(\bar{x}) + c(\operatorname{dist}(x, \bar{x}))$$

where $c : 0 \in D \subseteq [0, \infty) \to [0, \infty)$, $c(0) = 0$, and $c(a_n) \to 0 \Rightarrow a_n \to 0$. When f is convex, the forcing function c can be chosen convex, [19, Proposition 10.1.9]. Chapter III of [16] also contains several results on generic well-posedness. For further information on strong minimizers, we refer readers to [12, 14, 19].

Definition 2.20. *Following Rockafellar and Wets (see [27, p. 90]), we will call a function $f \in \Gamma_0(\mathbb{R}^n)$* level-coercive *if*

$$\liminf_{\|x\| \to \infty} \frac{f(x)}{\|x\|} > 0,$$

and coercive *if*

$$\liminf_{\|x\| \to \infty} \frac{f(x)}{\|x\|} = \infty.$$

Definition 2.21. *A function $f \in \Gamma_0(\mathbb{R}^n)$ is σ-*strongly convex *if there exists a constant $\sigma > 0$ such that $f - \frac{\sigma}{2}\|\cdot\|^2$ is convex. Equivalently, f is σ-strongly convex if there exists $\sigma > 0$ such that for all $\lambda \in (0, 1)$ and for all $x, y \in \mathbb{R}^n$,*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{\sigma}{2}\lambda(1 - \lambda)\|x - y\|^2.$$

Fact 2.22. ([10, Exercise 21 p. 83], [27, Theorem 11.8]) *Let $f \in \Gamma_0(\mathbb{R}^n)$. Then*

(i) *f is level-coercive if and only if* $0 \in \operatorname{int} \operatorname{dom} f^*$, *and*

(ii) *f is coercive if and only if* $\operatorname{dom} f^* = \mathbb{R}^n$.

Lemma 2.23. *The function $f \in \Gamma_0(\mathbb{R}^n)$ is σ-strongly convex if and only if $e_1 f$ is $\frac{\sigma}{\sigma+1}$-strongly convex.*

Proof. By [27, Proposition 12.60], f is σ-strongly convex if and only if ∇f^* is $\frac{1}{\sigma}$-Lipschitz for some $\sigma > 0$. Now

$$\begin{aligned} (e_1 f)^* &= f^* + \frac{1}{2}\|\cdot\|^2, \text{ and} \\ \nabla(e_1 f)^* &= \nabla f^* + \operatorname{Id}. \end{aligned}$$

Suppose that f is σ-strongly convex. Since ∇f^* is $\frac{1}{\sigma}$-Lipschitz, we have that $\nabla f^* + \operatorname{Id}$ is $\left(1 + \frac{1}{\sigma}\right)$-Lipschitz. Hence, $\nabla(e_1 f)^*$ is $\frac{\sigma+1}{\sigma}$-Lipschitz. Then $e_1 f$ is $\frac{\sigma}{\sigma+1}$-strongly convex, and we have proved one direction of the lemma. Working backwards with the same argument, the other direction is proved as well. $\square$

Lemma 2.24. *Let $f \in \Gamma_0(\mathbb{R}^n)$. Then*

(i) *f is level-coercive if and only if $e_1 f$ is level-coercive, and*

(ii) *f is coercive if and only if $e_1 f$ is coercive.*

Proof. Since $(e_1 f)^* = f^* + \frac{1}{2}\|\cdot\|^2$, we have $\operatorname{dom}(e_1 f)^* = \operatorname{dom} f^*$. The result follows from Fact 2.22. □

Lemma 2.25. *Let $f \in \Gamma_0(\mathbb{R}^n)$ be strongly convex. Then f is coercive.*

Proof. Since f is strongly convex, f can be written as $g + \frac{\sigma}{2}\|\cdot\|^2$ for some $g \in \Gamma_0(\mathbb{R}^n)$ and $\sigma > 0$. Since g is convex, g is bounded below by a hyperplane. That is, there exist $\tilde{x} \in \mathbb{R}^n$ and $r \in \mathbb{R}$ such that

$$g(x) \geq \langle \tilde{x}, x \rangle + r \text{ for all } x \in \mathbb{R}^n .$$

Hence,

$$f(x) \geq \langle \tilde{x}, x \rangle + r + \frac{\sigma}{2}\|x\|^2 \text{ for all } x \in \mathbb{R}^n .$$

This gives us that

$$\liminf_{\|x\|\to\infty} \frac{f(x)}{\|x\|} = \infty.$$

□

Lemma 2.26. *Let $f : \Gamma_0(\mathbb{R}^n) \to \overline{\mathbb{R}}$ be strongly convex. Then the (unique) minimizer of f is a strong minimizer.*

Proof. Let $f(x_k) \to \inf_x f(x)$. Since f is coercive by Lemma 2.25, $\{x_k\}_{k=1}^{\infty}$ is bounded. By the Bolzano-Weierstrass Theorem, $\{x_k\}_{k=1}^{\infty}$ has a convergent subsequence $x_{k_j} \to \bar{x}$. Since f is lsc, we have that $\liminf_{k\to\infty} f(x_k) \geq f(\bar{x})$. Hence,

$$\inf_x f(x) \leq f(\bar{x}) \leq \inf_x f(x).$$

Therefore, $f(\bar{x}) = \inf_x f(x)$. Since strong convexity implies strict convexity, $\operatorname{argmin} f(x) = \{\bar{x}\}$ is unique. As every subsequence of $\{x_k\}_{k=1}^{\infty}$ converges to the same limit $\bar{x}$, we conclude that $x_k \to \bar{x}$. □

Remark 2.27. *See [32, Proposition 3.5.8] for a stronger and more general result regarding Lemmas 2.25 and 2.26.*

In view of [32, Corollary 3.5.11(iii)], when f is strongly convex with strong minimizer $\bar{x}$, taking $(\bar{x}, 0) \in \operatorname{gr} \partial f$ one has the existence of $c > 0$ such that

$$f(x) \geq f(\bar{x}) + \frac{c}{2}\|x - \bar{x}\|^2.$$

From this, we have

$$\|x_k - \bar{x}\| \leq \sqrt{\frac{2}{c}[f(x_k) - f(\bar{x})]} \text{ for all } k \in \mathbb{N}\,.$$

This ensures a bound on the rate of convergence of the minimizing sequence to the minimizer in terms of function values.

Note that a convex function can be coercive, but fail to be strongly convex. Consider the following example.

Example 2.28. *Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$. The function f is coercive and attains a strong minimum at $\bar{x} = 0$, but is not strongly convex.*

Proof. It is clear that f is coercive. By definition, f is strongly convex if and only if there exists $\sigma > 0$ such that $g(x) := x^4 - \frac{\sigma}{2}x^2$ is convex. Since g is a differentiable, univariate function, we know it is convex if and only if its second derivative is nonnegative for all $x \in \mathbb{R}$. Since $g''(x) = 12x^2 - \sigma$ is clearly not nonnegative for any fixed $\sigma > 0$ and all $x \in \mathbb{R}$, we have that g is not convex. Therefore, f is not strongly convex. Clearly, zero is the minimum and minimizer of f. Let $\{x_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ be such that $f(x_k) \to f(0) = 0$. Then $\lim\limits_{n\to\infty} x_k^4 = 0$ implies $\lim\limits_{n\to\infty} x_k = 0$. Therefore, f attains a strong minimum. □

The following additional counterexample is thanks to one of the referees.

Example 2.29. *Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) := |x|^p$, where $p > 1$. Then f is coercive for all such p, but f is strongly convex if and only if $p = 2$.*

Proof. It is elementary to show that f is coercive. Let $\sigma > 0$, and define $g(x) := f(x) - \frac{\sigma}{2}x^2$. Then $g''(x) = p(p-1)x^{p-2} - \sigma$. In order to conclude that f is strongly convex, we must have

$$g''(x) \geq 0 \text{ for all } x \in \mathbb{R}\,. \tag{2.1}$$

However, if $p > 2$ then inequality (2.1) fails for $|x|$ small, specifically for $|x| < \left(\frac{\sigma}{p(p-1)}\right)^{\frac{1}{p-2}}$, and if $1 < p < 2$ then inequality (2.1) fails for $|x|$ large. Only for $p = 2$ can we choose σ such that inequality (2.1) is true for all x. □

3 A complete metric space using Moreau envelopes

The principal tool we use is the Baire category theorem. To this end, we need a Baire space. In this section, we establish a complete metric space whose distance function makes use of the Moreau envelope. This metric has been used by Attouch and Wets in [3, page 38]. The distances used in the next section refer to the metric established here. In Beer [7, p. 241], it is left as an exercise to show that the space defined in this section is a complete metric space. For the purpose of self-containment, we include a full proof here.

We begin with some properties on the Moreau envelope. Set

$$e_1(\Gamma_0(\mathbb{R}^n)) := \{e_1 f : \ f \in \Gamma_0(\mathbb{R}^n)\}.$$

Theorem 3.1. *The set* $e_1(\Gamma_0(\mathbb{R}^n))$ *is a convex set in* $\Gamma_0(\mathbb{R}^n)$.

Proof. Let $f_1, f_2 \in \Gamma_0(\mathbb{R}^n)$, $\lambda \in [0,1]$. Then $e_1 f_1, e_1 f_2 \in e_1(\Gamma_0(\mathbb{R}^n))$. We need to show that $\lambda e_1 f_1 + (1-\lambda) e_1 f_2 \in e_1(\Gamma_0(\mathbb{R}^n))$. By [6, Theorem 6.2] with $\mu = 1$ and $n = 2$, we have that $\lambda e_1 f_1 + (1-\lambda) e_1 f_2$ is the Moreau envelope of the proximal average function $P_1(f, \lambda)$. By [6, Corollary 5.2], we have that $P_1(f,\lambda) \in \Gamma_0(\mathbb{R}^n)$. Hence, $e_1 P_1(f,\lambda) \in e_1(\Gamma_0(\mathbb{R}^n))$, and we conclude that $e_1(\Gamma_0(\mathbb{R}^n))$ is a convex set. □

On $e_1(\Gamma_0(\mathbb{R}^n))$, define a metric by

$$\tilde{d}(\tilde{f}, \tilde{g}) := \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\|\tilde{f} - \tilde{g}\|_i}{1 + \|\tilde{f} - \tilde{g}\|_i}, \tag{3.1}$$

where $\|\tilde{f} - \tilde{g}\|_i := \sup_{\|x\| \leq i} |\tilde{f}(x) - \tilde{g}(x)|$ and $\tilde{f}, \tilde{g} \in e_1(\Gamma_0(\mathbb{R}^n))$.

Note that a sequence of functions in $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$ converges if and only if the sequence converges uniformly on bounded sets, if and only if the sequence converges pointwise on $\mathbb{R}^n$. The latter fails in infinite-dimensional space.

Theorem 3.2. *The metric space* $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$ *is complete.*

Proof. Consider a Cauchy sequence $\{e_1 f_k\}_{k \in \mathbb{N}}$ in $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$ where $f_k \in \Gamma_0(\mathbb{R}^n)$ for $k \in \mathbb{N}$. Since $f_k \in \Gamma_0(\mathbb{R}^n)$, by Fact 2.10, $e_1 f_k$ is continuous and differentiable on $\mathbb{R}^n$. Then $e_1 f_k \xrightarrow{p} g$ where $g : \mathbb{R}^n \to \mathbb{R}$ is continuous and convex. Our objective is to prove that g is in fact the Moreau envelope of a proper, lsc, convex function. Since $(e_1 f_k)_{k \in \mathbb{N}}$ and g are convex and full-domain, by [27, Theorem 7.17] we have that $e_1 f_k \xrightarrow{e} g$, and $e_1 f_k \xrightarrow{u} g$ on bounded sets of $\mathbb{R}^n$. By [27, Theorem 11.34], we have that $(e_1 f_k)^* \xrightarrow{e} g^*$, that is, $f_k^* + \frac{1}{2}\|\cdot\|^2 \xrightarrow{e} g^*$ and g^* is proper, lsc, and convex. Hence, $f_k^* \xrightarrow{e} g^* - \frac{1}{2}\|\cdot\|^2$ by [27, Exercise 7.8(a)]. As $\{f_k^*\}_{k \in \mathbb{N}}$ is a sequence of convex functions, by [27, Theorem 7.17] we have that $g^* - \frac{1}{2}\|\cdot\|^2$ is convex. Since $g^* - \frac{1}{2}\|\cdot\|^2$ is proper, lsc, and convex, there exists $h \in \Gamma_0(\mathbb{R}^n)$ such that $g^* - \frac{1}{2}\|\cdot\|^2 = h^*$. Applying [27, Theorem 11.34] again, we obtain $f_k \xrightarrow{e} h$. Finally, using Fact 2.15 or [27, Theorem 7.37] we see that $e_1 f_k \xrightarrow{p} e_1 h$, and $e_1 f_k \xrightarrow{u} e_1 h$ on bounded subsets of $\mathbb{R}^n$ as well. Therefore, $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$ is complete. [1] □

In view of Fact 2.15, we give the definition of the Attouch-Wets metric on $\Gamma_0(\mathbb{R}^n)$ as follows.

Definition 3.3 (Attouch-Wets metric)**.** *For* $f, g \in \Gamma_0(\mathbb{R}^n)$, *define the distance function* d :

$$d(f,g) := \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\|e_1 f - e_1 g\|_i}{1 + \|e_1 f - e_1 g\|_i}.$$

[1] A referee points out that this can also be obtained by proving that the σ-strong convexity property is preserved in the limit, using the characterization of the strong convexity in terms of the subdifferentials, and the corresponding graphical convergence of the subdifferentials.

In $\Gamma_0(\mathbb{R}^n)$, there are other metrics that induce the same topology as the Attouch-Wets metric; see Remark 3.6 below. In order to prove completeness of the space, we state the following lemma, whose simple proof is omitted.

Lemma 3.4. *Define* $a : [0, \infty) \to \mathbb{R}$, $a(t) := \frac{t}{1+t}$. *Then*

a) a *is an increasing function, and*

b) $t_1, t_2 \geq 0$ *implies that* $a(t_1 + t_2) \leq a(t_1) + a(t_2)$.

Proposition 3.5. *The space* $(\Gamma_0(\mathbb{R}^n), d)$ *where* d *is the metric defined in Definition 3.3, is a complete metric space.*

Proof. Items M1-M4 below show that $(\Gamma_0(\mathbb{R}^n), d)$ is a metric space, and item C that follows shows that it is complete.
M1: Since

$$\sum_{i=1}^{\infty} \frac{1}{2^i} = 1, \text{ and } 0 \leq \frac{\|e_1 f - e_1 g\|_i}{1 + \|e_1 f - e_1 g\|_i} < 1 \text{ for all } i,$$

we have that

$$\frac{1}{2^i} \geq \frac{1}{2^i} \frac{\|e_1 f - e_1 g\|_i}{1 + \|e_1 f - e_1 g\|_i} \text{ for all } i.$$

Then

$$0 \leq d(f, g) \leq 1 \text{ for all } f, g \in \Gamma_0(\mathbb{R}^n).$$

Hence, d is real-valued, finite, and non-negative.
M2: We have

$$\begin{aligned} d(f, g) = 0 &\Leftrightarrow \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\|e_1 f - e_1 g\|_i}{1 + \|e_1 f - e_1 g\|_i} = 0, \\ &\Leftrightarrow \|e_1 f - e_1 g\|_i = 0 \text{ for all } i, \\ &\Leftrightarrow e_1 f(x) - e_1 g(x) = 0 \text{ for all } x, \\ &\Leftrightarrow e_1 f = e_1 g, \\ &\Leftrightarrow f = g \text{ [27, Corollary 3.36]}. \end{aligned}$$

Hence, $d(f, g) = 0$ if and only if $f = g$.
M3: The fact that $d(f, g) = d(g, f)$ is trivial.
M4: By the triangle inequality,

$$\|e_1 f - e_1 g\|_i \leq \|e_1 f - e_1 h\|_i + \|e_1 h - e_1 g\|_i \text{ for all } f, g, h \in \Gamma_0(\mathbb{R}^n).$$

By applying Lemma 3.4(a), we have

$$\frac{\|e_1 f - e_1 g\|_i}{1 + \|e_1 f - e_1 g\|_i} \leq \frac{\|e_1 f - e_1 h\|_i + \|e_1 h - e_1 g\|_i}{1 + \|e_1 f - e_1 h\|_i + \|e_1 h - e_1 g\|_i}.$$

Then we apply Lemma 3.4(b) with $t_1 = \|e_1 f - e_1 h\|_i$ and $t_2 = \|e_1 h - e_1 g\|_i$, and we have

$$\frac{\|e_1 f - e_1 g\|_i}{1 + \|e_1 f - e_1 g\|_i} \le \frac{\|e_1 f - e_1 h\|_i}{1 + \|e_1 f - e_1 h\|_i} + \frac{\|e_1 h - e_1 g\|_i}{1 + \|e_1 h - e_1 g\|_i}.$$

Multiplying both sides by $\frac{1}{2^i}$ and taking the summation over i, we obtain the distance functions, which yields $d(f,g) \le d(f,h) + d(h,g)$ for all $f, g, h \in \Gamma_0(\mathbb{R}^n)$.

C: Let $\{f_k\}_{k=1}^\infty$ be a Cauchy sequence in $(\Gamma_0(\mathbb{R}^n), d)$. Then for each $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that $d(f_j, f_k) < \varepsilon$ for all $j, k \ge N_\varepsilon$. Fix $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that

$$\sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\|e_1 f_j - e_1 f_k\|_i}{1 + \|e_1 f_j - e_1 f_k\|_i} < \varepsilon \text{ for all } j, k \ge N.$$

Then for any $i \in \mathbb{N}$ fixed, we have $\frac{\|e_1 f_j - e_1 f_k\|_i}{1+\|e_1 f_j - e_1 f_k\|_i} < 2^i \varepsilon$, so that $\|e_1 f_j - e_1 f_k\|_i < \frac{2^i \varepsilon}{1 - 2^i \varepsilon} =: \hat{\varepsilon} > 0$, for all $j, k \ge N$. Notice that $\hat{\varepsilon} \searrow 0$ as $\varepsilon \searrow 0$. This gives us that $\{e_1 f_k\}_{k=1}^\infty$ is a Cauchy sequence on $\mathbb{B}_i(0)$ for each $i \in \mathbb{N}$, so that $e_1 f_k \xrightarrow{p} g$ for some finite-valued convex function g. By the same arguments as in the proof of Theorem 3.2, we know that $g = e_1 h$ with $h \in \Gamma_0(\mathbb{R}^n)$. Hence $f_k \to h$ in terms of metric d. Therefore, $(\Gamma_0(\mathbb{R}^n), d)$ is closed, and is a complete metric space. $\square$

Remark 3.6. *(1). This result can also be reached via [2, Theorem 2.1], using the ρ-Hausdorff distance for epigraphs; and this result is also mentioned as an exercise in [7, p. 241]. We refer the reader to [7, Chapter 7] for more details on the Attouch-Wets topology for convex functions.*

(2). When X is infinite dimensional, on the set of proper lower semicontinuous convex function $\Gamma_0(X)$, a variety different topologies show up, such as Kuratowski-Painlevé convergence, Attouch-Wets convergence, Mosco convergence, Choquet-Wijsman convergence, etc. See Beer [7], Lucchetti [19], Borwein-Vanderwerff [11], Attouch [1]. However, when X is finite dimensional, all these convergences coincide [11, Theorem 6.2.13].

On the set of Fenchel conjugates

$$(\Gamma_0(\mathbb{R}^n))^* := \{f^* : \ f \in \Gamma_0(\mathbb{R}^n)\}$$

define a metric by $\hat{d}(f^*, g^*) := d(f^*, g^*)$ for $f^*, g^* \in (\Gamma_0(\mathbb{R}^n))^*$. Observe that $\Gamma_0(\mathbb{R}^n) = (\Gamma_0(\mathbb{R}^n))^*$.

Corollary 3.7. *Consider two metric spaces $(\Gamma_0(\mathbb{R}^n), d)$ and $((\Gamma_0(\mathbb{R}^n))^*, \hat{d})$. Define*

$$T : (\Gamma_0(\mathbb{R}^n), d) \to ((\Gamma_0(\mathbb{R}^n))^*, \hat{d}) : f \mapsto f^*.$$

Then T is a bijective isometry. Consequently, $(\Gamma_0(\mathbb{R}^n), d)$ and $((\Gamma_0(\mathbb{R}^n))^*, \hat{d})$ are isometric.

Proof. Clearly T is onto. Also, T is injective because of the Fenchel-Moreau Theorem [5, Theorem 13.32] or [26, Corollary 12.2.1]. To see this, let $Tf = Tg$. Then $f^* = g^*$, so $f = (f^*)^* = (g^*)^* = g$. It remains to show that T is an isometry:

$$(\forall f, g \in \Gamma_0(\mathbb{R}^n)) \quad d(f,g) = d(f^*, g^*) = \hat{d}(Tf, Tg).$$

To see this, using Lemma 2.18 we have

$$\begin{aligned}
d(f^*, g^*) &= \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\sup\limits_{\|x\|\le i} |e_1 f^*(x) - e_1 g^*(x)|}{1 + \sup\limits_{\|x\|\le i} |e_1 f^*(x) - e_1 g^*(x)|} \\
&= \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\sup\limits_{\|x\|\le i} \left|\frac{1}{2}\|x\|^2 - e_1 f(x) - \frac{1}{2}\|x\|^2 + e_1 g(x)\right|}{1 + \sup\limits_{\|x\|\le i} \left|\frac{1}{2}\|x\|^2 - e_1 f(x) - \frac{1}{2}\|x\|^2 + e_1 g(x)\right|} \\
&= \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\sup\limits_{\|x\|\le i} |e_1 g(x) - e_1 f(x)|}{1 + \sup\limits_{\|x\|\le i} |e_1 g(x) - e_1 f(x)|} \\
&= d(f, g).
\end{aligned}$$

□

By Theorem 3.2, $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$ is a complete metric space.

Corollary 3.8. *Consider two metric spaces* $(\Gamma_0(\mathbb{R}^n), d)$ *and* $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$*. Define*

$$T : \Gamma_0(\mathbb{R}^n) \to e_1(\Gamma_0(\mathbb{R}^n)) : f \mapsto e_1 f.$$

Then T is a bijective isometry, so $(\Gamma_0(\mathbb{R}^n), d)$ *and* $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$ *are isometric.*

4 Baire category results

This section is devoted to the main work of this paper. Ultimately, we show that the set of strongly convex functions is a meager (Baire category one) set, while the set of convex functions that attain a strong minimum is a generic (Baire category two) set.

4.1 Characterizations of the strong minimizer

The first proposition describes the relationship between a function and its Moreau envelope, pertaining to the strong minimum. Several more results regarding strong minima follow.

Proposition 4.1. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then f attains a strong minimum at $\bar{x}$ if and only if $e_1 f$ attains a strong minimum at $\bar{x}$.*

Proof. $(\Rightarrow)$ Assume that f attains a strong minimum at $\bar{x}$. Then

$$\min_x f(x) = \min_x e_1 f(x) = f(\bar{x}) = e_1 f(\bar{x}).$$

Let $\{x_k\}$ be such that $e_1 f(x_k) \to e_1 f(\bar{x})$. We need to show that $x_k \to \bar{x}$. Since

$$e_1 f(x_k) = f(v_k) + \frac{1}{2}\|v_k - x_k\|^2$$

for some v_k, and $f(v_k) \geq f(\bar{x})$, we have

$$0 \leq \frac{1}{2}\|x_k - v_k\|^2 + f(v_k) - f(\bar{x}) = e_1 f(x_k) - e_1 f(\bar{x}) \to 0. \tag{4.1}$$

Since both $\frac{1}{2}\|x_k - v_k\|^2 \geq 0$ and $f(v_k) - f(\bar{x}) \geq 0$, equation (4.1) tells us that $x_k - v_k \to 0$ and $f(v_k) \to f(\bar{x})$. Since $\bar{x}$ is the strong minimizer of f, we have $v_k \to \bar{x}$. Therefore, $x_k \to \bar{x}$, and $e_1 f$ attains a strong minimum at $\bar{x}$.
($\Leftarrow$) Assume that $e_1 f$ attains a strong minimum at $\bar{x}$, $e_1 f(\bar{x}) = \min e_1 f$. Then $e_1 f(x_k) \to e_1 f(\bar{x})$ implies that $x_k \to \bar{x}$. Let $f(x_k) \to f(\bar{x})$. We have

$$f(\bar{x}) \leq e_1 f(\bar{x}) \leq e_1 f(x_k) \leq f(x_k).$$

Since $f(x_k) \to f(\bar{x})$, we obtain

$$e_1 f(x_k) \to f(\bar{x}) = e_1 f(\bar{x}).$$

Therefore, $x_k \to \bar{x}$, and f attains a strong minimum at $\bar{x}$. $\square$

Proposition 4.2. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ have a strong minimizer $\bar{x}$. Then for all $m \in \mathbb{N}$,*

$$\inf_{\|x-\bar{x}\| \geq \frac{1}{m}} f(x) > f(\bar{x}).$$

Proof. This is clear by the definition of strong minimizer. Indeed, suppose that there exists $m \in \mathbb{N}$ such that $\inf_{\|x-\bar{x}\| \geq \frac{1}{m}} f(x) = f(\bar{x})$. Then there exists a sequence $\{x_k\}_{k=1}^{\infty}$ with $\|x_k - \bar{x}\| \geq \frac{1}{m}$ and $\lim_{k\to\infty} f(x_k) = f(\bar{x})$. Since $\bar{x}$ is the strong minimizer of f, we have $x_k \to \bar{x}$, a contradiction. $\square$

Corollary 4.3. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ have a strong minimizer $\bar{x}$. Then for all $m \in \mathbb{N}$,*

$$\inf_{\|x-\bar{x}\| \geq \frac{1}{m}} e_1 f(x) > e_1 f(\bar{x}).$$

Proof. This follows directly from Propositions 4.1 and 4.2. $\square$

The next result describes a distinguished property of convex functions defined in $\mathbb{R}^n$.

Theorem 4.4. *Let $f \in \Gamma_0(\mathbb{R}^n)$. Then f has a strong minimizer if and only if f has a unique minimizer.*

Proof. $(\Rightarrow)$ By definition, if f has a strong minimizer, then that minimizer is unique.
$(\Leftarrow)$ Suppose f has a unique minimizer $\bar{x}$. Because $f \in \Gamma_0(\mathbb{R}^n)$, by [26, Theorem 8.7], all sublevel-sets $\{x : f(x) \le \alpha\}$, for any $\alpha \ge f(\bar{x})$, have the same recession cone. Since the recession cone of $\{x : f(x) \le f(\bar{x})\} = \{\bar{x}\}$ is 0, each sublevel set of f is bounded. Since f has a unique minimizer $\bar{x}$ and all sublevel sets of f are bounded, we have that $\bar{x}$ is in fact a strong minimizer. Indeed, this follows by applying [11, Fact 4.4.8], [11, Theorem 4.4.10], and [11, Theorem 5.23(e)(c)] in $\mathbb{R}^n$ because $\partial f^*(0) = \operatorname{argmin} f = \{\bar{x}\}$. □

Remark 4.5. *See also [11, Exercise 5.2.1 p. 234].*

Remark 4.6. *A more self-contained proof of Theorem 4.4, more in the style of that of Lemma 4.20, is provided below.*

Proof. Let $f(x_k) \to f(\bar{x})$. We need to show that $x_k \to \bar{x}$. Suppose the contrary, $x_k \not\to \bar{x}$. Taking a subsequence if necessary, we can assume that there exists $m \in \mathbb{N}$ such that

$$\|x_k - \bar{x}\| \ge \frac{1}{m} \text{ for all } k \in \mathbb{N}\,.$$

Select y_k such that

$$y_k \in [x_k, \bar{x}] \cap \left\{x : m \ge \|x - \bar{x}\| \ge \frac{1}{m}\right\}.$$

Then for all k, we have that $\|y_k\| \le m$, and that there exists $0 \le \lambda_k \le 1$ such that

$$y_k = \lambda_k x_k + (1 - \lambda_k)\bar{x}.$$

By the convexity of f, we have that for all k,

$$f(y_k) \le \lambda_k f(x_k) + (1 - \lambda_k) f(\bar{x}).$$

As $\{y_k\}$ is bounded, λ_k is bounded. Taking convergent subsequences if necessary, we may assume that

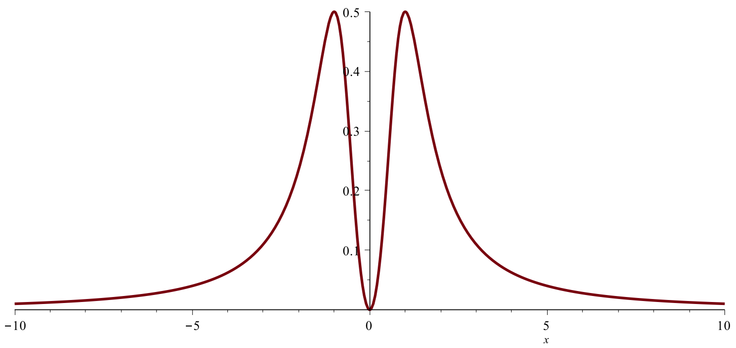
$$y_k \to \bar{y}, \qquad \text{and} \qquad \lambda_k \to \bar{\lambda}.$$

Then

$$\begin{aligned} f(\bar{y}) &\le \liminf_{k\to\infty} f(y_k) \\ &\le \bar{\lambda} f(\bar{x}) + (1 - \bar{\lambda}) f(\bar{x}) \\ &= f(\bar{x}). \end{aligned}$$

As $f(\bar{y}) \ge f(\bar{x})$, we have $f(\bar{y}) = f(\bar{x})$. Since $\|\bar{y} - \bar{x}\| \ge \frac{1}{m}$, we have $\bar{y} \ne \bar{x}$, which contradicts the fact that $\bar{x}$ is the unique minimizer of f. □

Example 4.7. *Theorem 4.4 can fail when the function is nonconvex. Consider the continuous but nonconvex function* $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{x^2}{x^4+1}$.



The function has a unique minimizer $\bar{x} = 0$, *but the minimizer is not strong, as any sequence* $\{x_k\}$ *that tends to* $\pm\infty$ *gives a sequence of function values that tends to* $f(\bar{x})$.

Example 4.8. *Theorem 4.4 can also fail in infinite-dimensional space. Consider the continuous, convex function*

$$f : l^2 \to \overline{\mathbb{R}} : x \mapsto f(x) := \sum_{k=1}^{\infty} \frac{1}{k} x_k^2.$$

This function has a unique minimizer $\bar{x} = 0$, *but* $\bar{x}$ *is not a strong minimizer because* $f(e_k) = \frac{1}{k} \to 0$ *and* $e_k \not\to 0$. *Here,* $e_k = (0, \ldots, 0, 1, 0, \ldots)$, *where the* 1 *is in the* k*th position. See also [7, p. 268, Exercise 2].*

Using Proposition 4.2 and Corollary 4.3, we can now single out two sets in $\Gamma_0(\mathbb{R}^n)$ which are very important for our later proofs.

Definition 4.9. *For any* $m \in \mathbb{N}$, *define the sets* U_m *and* E_m *as follows:*

$$U_m := \left\{ f \in \Gamma_0(\mathbb{R}^n) : \text{ there exists } z \in \mathbb{R}^n \text{ such that } \inf_{\|x-z\| \geq \frac{1}{m}} f(x) - f(z) > 0 \right\},$$

$$E_m := \left\{ f \in \Gamma_0(\mathbb{R}^n) : \text{ there exists } z \in \mathbb{R}^n \text{ such that } \inf_{\|x-z\| \geq \frac{1}{m}} e_1 f(x) - e_1 f(z) > 0 \right\}.$$

Proposition 4.10. *Let* $f \in \bigcap_{m \in \mathbb{N}} U_m$. *Then* f *attains a strong minimum on* $\mathbb{R}^n$.

Proof. The proof follows the method of [14, Theorem II.1]. Since $f \in \bigcap_{m \in \mathbb{N}} U_m$, we have that for each $m \in \mathbb{N}$ there exists $z_m \in \mathbb{R}^n$ such that

$$f(z_m) < \inf_{\|x-z_m\| \geq \frac{1}{m}} f(x).$$

Suppose that $\|z_p - z_m\| \geq \frac{1}{m}$ for some $p > m$. By the definition of z_m, we have

$$f(z_p) > f(z_m). \tag{4.2}$$

Since $\|z_m - z_p\| \geq \frac{1}{m} > \frac{1}{p}$, we have

$$f(z_m) > f(z_p)$$

by the definition of z_p. This contradicts inequality (4.2). Thus, $\|z_p - z_m\| < \frac{1}{m}$ for each $p > m$. This gives us that $\{z_m\}_{m=1}^{\infty}$ is a Cauchy sequence that converges to some $\bar{x} \in \mathbb{R}^n$. It remains to be shown that $\bar{x}$ is the strong minimizer of f. Since f is lsc, we have

$$\begin{aligned} f(\bar{x}) &\leq \liminf_{m\to\infty} f(z_m) \\ &\leq \liminf_{m\to\infty} \left(\inf_{\|x-z_m\|\geq \frac{1}{m}} f(x) \right) \\ &\leq \inf_{x\in\mathbb{R}^n\setminus\{\bar{x}\}} f(x). \end{aligned}$$

Let $\{y_k\}_{k=1}^{\infty} \subseteq \mathbb{R}^n$ be such that $f(y_k) \to f(\bar{x})$, and suppose that $y_k \not\to \bar{x}$. Dropping to a subsequence if necessary, there exists $\varepsilon > 0$ such that $\|y_k - \bar{x}\| \geq \varepsilon$ for all k. Thus, there exists $p \in \mathbb{N}$ such that $\|y_k - z_p\| \geq \frac{1}{p}$ for all $k \in \mathbb{N}$. Hence,

$$f(\bar{x}) \leq f(z_p) < \inf_{\|x-z_p\|\geq \frac{1}{p}} f(x) \leq f(y_k)$$

for all $k \in \mathbb{N}$, a contradiction to the fact that $f(y_k) \to f(\bar{x})$. Therefore, $\bar{x}$ is the strong minimizer of f. □

Theorem 4.11. *Let $f \in \bigcap_{m\in\mathbb{N}} E_m$. Then $e_1 f$ attains a strong minimum on $\mathbb{R}^n$, so f attains a strong minimum on $\mathbb{R}^n$.*

Proof. Applying Proposition 4.10, for each $f \in \bigcap_{m\in\mathbb{N}} E_m$, $e_1 f$ has a strong minimizer on $\mathbb{R}^n$. By Proposition 4.1, each corresponding f has the same corresponding strong minimizer. □

4.2 The set of strongly convex functions is dense, but first category

Next, we turn our attention to the set of strongly convex functions. The objectives here are to show that the set is contained in both U_m and E_m, dense in $(\Gamma_0(\mathbb{R}^n), d)$, and meager in $(\Gamma_0(\mathbb{R}^n), d)$.

Theorem 4.12. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be strongly convex. Then $f \in U_m$ and $f \in E_m$ for all $m \in \mathbb{N}$.*

Proof. Since f is strongly convex, f has a unique minimizer z. By Lemma 2.26, z is a strong minimizer, so that for any sequence $\{x_k\}$ such that $f(x_k) \to f(\bar{x})$, we must have $x_k \to \bar{x}$. We want to show that

$$\inf_{\|x-z\|\geq \frac{1}{m}} f(x) - f(z) > 0. \tag{4.3}$$

For any $m \in \mathbb{N}$, equation (4.3) is true by Proposition 4.2. Therefore, $f \in U_m$ for all $m \in \mathbb{N}$. By Lemma 2.23, $e_1 f$ is strongly convex. Therefore, by the same reasoning as above, $f \in E_m$ for all $m \in \mathbb{N}$. □

We will need the following characterizations of strongly convex functions in later proofs. Note that (i)⇒(iii) has been done by Rockafellar [25].

Lemma 4.13. *Let $f \in \Gamma_0(\mathbb{R}^n)$. The following are equivalent:*

(i) *f is strongly convex.*

(ii) *$\operatorname{Prox}_f^1 = k \operatorname{Prox}_g^1$ for some $0 \le k < 1$ and $g \in \Gamma_0(\mathbb{R}^n)$.*

(iii) *$\operatorname{Prox}_f^1 = kN$ for some $0 \le k < 1$ and $N : \mathbb{R}^n \to \mathbb{R}^n$ nonexpansive.*

Proof. (i)⇒(ii): Assume that f is strongly convex. Then $f = g + \sigma q$ where $g \in \Gamma_0(\mathbb{R}^n)$, $q = \frac{1}{2}\|\cdot\|^2$, and $\sigma > 0$. We have

$$\operatorname{Prox}_f^1 = ((1+\sigma)\operatorname{Id} + \partial g)^{-1} = \left((1+\sigma)\big(\operatorname{Id} + \frac{\partial g}{1+\sigma}\big)\right)^{-1} \tag{4.4}$$

$$= \left(\operatorname{Id} + \frac{\partial g}{1+\sigma}\right)^{-1}\left(\frac{\operatorname{Id}}{1+\sigma}\right). \tag{4.5}$$

Define $\tilde{g}(x) = (1+\sigma)g(x/(1+\sigma))$. Then $\tilde{g} \in \Gamma_0(\mathbb{R}^n)$, $\partial \tilde{g} = \partial g \circ \big(\frac{\operatorname{Id}}{1+\sigma}\big)$, so

$$\operatorname{Prox}_{\tilde{g}}^1 = \left(\operatorname{Id} + \partial g \circ \left(\frac{\operatorname{Id}}{1+\sigma}\right)\right)^{-1} = \left((1+\sigma)\left(\operatorname{Id} + \frac{\partial g}{1+\sigma}\right) \circ \left(\frac{\operatorname{Id}}{1+\sigma}\right)\right)^{-1} \tag{4.6}$$

$$= (1+\sigma)\left(1 + \frac{\partial g}{1+\sigma}\right)^{-1} \circ \left(\frac{\operatorname{Id}}{1+\sigma}\right) \tag{4.7}$$

$$= (1+\sigma)\operatorname{Prox}_f^1. \tag{4.8}$$

Therefore, $\operatorname{Prox}_f^1 = \frac{1}{1+\sigma}\operatorname{Prox}_{\tilde{g}}^1$.

(ii)⇒(i): Assume $\operatorname{Prox}_f^1 = k \operatorname{Prox}_g^1$ for some $0 \le k < 1$ and $g \in \Gamma_0(\mathbb{R}^n)$. If $k = 0$, then $f = \iota_{\{0\}}$, and f is obviously strongly convex. Let us assume $0 < k < 1$. The assumption $(\operatorname{Id} + \partial f)^{-1} = k(\operatorname{Id} + \partial g)^{-1}$ gives $\operatorname{Id} + \partial f = (\operatorname{Id} + \partial g) \circ (\operatorname{Id}/k) = \operatorname{Id}/k + \partial g \circ (\operatorname{Id}/k)$, so

$$\partial f = (1/k - 1)\operatorname{Id} + \partial g(\operatorname{Id}/k).$$

Since $1/k > 1$ and $\partial g \circ (\operatorname{Id}/k)$ is monotone, we have that ∂f is strongly monotone, which implies that f is strongly convex.

(ii)⇒(iii): This is clear because Prox_g^1 is nonexpansive, see, e.g., [5, Proposition 12.27].

(iii)⇒(ii): Assume $\operatorname{Prox}_f^1 = kN$ where $0 \le k < 1$ and N is nonexpansive. If $k = 0$, then $\operatorname{Prox}_f^1 = 0 = 0 \cdot 0$, so (ii) holds because $\operatorname{Prox}_{\iota_{\{0\}}}^1 = 0$. If $0 < k < 1$, then $N = 1/k \operatorname{Prox}_f^1$. As

$$\operatorname{Prox}_f^1 = (\operatorname{Id} + \partial f)^{-1} = \nabla(q + f)^* = \nabla e_1(f^*),$$

we have $N = \nabla(e_1(f^*)/k)$. This means that N is nonexpansive and the gradient of a differentiable convex function. By the Baillon-Haddad theorem [4] or [5, Corollary 18.16], $N = \operatorname{Prox}_g^1$ for some $g \in \Gamma_0(\mathbb{R}^n)$. Therefore, $\operatorname{Prox}_f^1 = k \operatorname{Prox}_g^1$, i.e., (ii) holds true. $\square$

Theorem 4.14. *The set of strongly convex functions is dense in* $(\Gamma_0(\mathbb{R}^n), d)$. *Equivalently, the set of strongly convex functions is dense in* $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$.

Proof. Let $0 < \varepsilon < 1$ and $f \in \Gamma_0(\mathbb{R}^n)$. It will suffice to find $h \in \Gamma_0(\mathbb{R}^n)$ such that h is strongly convex and $d(h, f) < \varepsilon$. For $0 < \sigma < 1$, define $g \in \Gamma_0(\mathbb{R}^n)$ by way of the proximal mapping:

$$\operatorname{Prox}_g^1 := (1-\sigma)\operatorname{Prox}_f^1 = (1-\sigma)\operatorname{Prox}_f^1 + \sigma \operatorname{Prox}_{\iota_{\{0\}}}^1 .$$

Such a $g \in \Gamma_0(\mathbb{R}^n)$ does exists because g is the proximal average of f and $\iota_{\{0\}}$ by [6], and g is strongly convex because of Lemma 4.13. Define $h \in \Gamma_0(\mathbb{R}^n)$ by

$$h := g - e_1 g(0) + e_1 f(0).$$

Indeed, some calculations give

$$h = (1-\sigma) f\left(\frac{\cdot}{1-\sigma}\right) + \frac{\sigma}{1-\sigma} q + \sigma e_1 f(0). \tag{4.9}$$

Then $e_1 h = e_1 g - e_1 g(0) + e_1 f(0)$, so that

$$e_1 h(0) = e_1 f(0), \tag{4.10}$$

and $\operatorname{Prox}_h^1 = \operatorname{Prox}_g^1$. Fix N large enough that $\sum_{i=N}^{\infty} \frac{1}{2^i} < \frac{\varepsilon}{2}$. Then

$$\sum_{i=N}^{\infty} \frac{1}{2^i} \frac{\|e_1 f - e_1 g\|_i}{1 + \|e_1 f - e_1 g\|_i} \leq \sum_{i=N}^{\infty} \frac{1}{2^i} < \frac{\varepsilon}{2}. \tag{4.11}$$

Choose σ such that

$$0 < \sigma < \frac{\varepsilon}{2-\varepsilon} \frac{1}{N(N + \|\operatorname{Prox}_f^1(0)\|)}. \tag{4.12}$$

This gives us that

$$\frac{\sigma N(N + \|\operatorname{Prox}_f^1(0)\|)}{1 + \sigma N(N + \|\operatorname{Prox}_f^1(0)\|)} < \frac{\varepsilon}{2}. \tag{4.13}$$

By equation (4.10) and the Mean Value Theorem, for some $c \in [x, 0]$ we have

$$\begin{aligned}
e_1 h(x) - e_1 f(x) &= e_1 h(x) - e_1 f(x) - (e_1 h(0) - e_1 f(0)) \\
&= \langle \nabla e_1 h(c) - \nabla e_1 f(c), x - 0 \rangle \\
&= \langle (\operatorname{Id} - \operatorname{Prox}_h^1)(c) - (\operatorname{Id} - \operatorname{Prox}_f^1)(c), x - 0 \rangle \\
&= \langle -\operatorname{Prox}_h^1(c) + \operatorname{Prox}_f^1(c), x - 0 \rangle \\
&= \langle -(1-\sigma)\operatorname{Prox}_f^1(c) + \operatorname{Prox}_f^1(c), x \rangle \\
&= \langle \sigma \operatorname{Prox}_f^1(c), x \rangle.
\end{aligned}$$

Using the triangle inequality, the Cauchy-Schwarz inequality, and the fact that Prox_f^1 is nonexpansive, we obtain

$$\begin{aligned}
|e_1h(x) - e_1f(x)| &\leq \sigma\| \operatorname{Prox}_f^1(c)\|\|x\| \\
&= \sigma\| \operatorname{Prox}_f^1(c) - \operatorname{Prox}_f^1(0) + \operatorname{Prox}_f^1(0)\|\|x\| \\
&\leq \sigma(\| \operatorname{Prox}_f^1(c) - \operatorname{Prox}_f^1(0)\| + \| \operatorname{Prox}_f^1(0)\|)\|x\| \\
&\leq \sigma(\|c\| + \| \operatorname{Prox}_f^1(0)\|)\|x\| \\
&\leq \sigma(\|x\| + \| \operatorname{Prox}_f^1(0)\|)\|x\| \\
&\leq \sigma N(N + \| \operatorname{Prox}_f^1(0)\|),
\end{aligned}$$

when $\|x\| \leq N$. Therefore, $\|e_1h - e_1f\|_N \leq \sigma N(N + \| \operatorname{Prox}_f^1(0)\|)$. Applying equation (4.13), this implies that

$$\frac{\|e_1f - e_1g\|_N}{1 + \|e_1f - e_1g\|_N} \leq \frac{\sigma N(N + \| \operatorname{Prox}_f^1(0)\|)}{1 + \sigma N(N + \| \operatorname{Prox}_f^1(0)\|)} < \frac{\varepsilon}{2}. \tag{4.14}$$

Now considering the first $N - 1$ terms of our d function, we have

$$\begin{aligned}
\sum_{i=1}^{N-1} \frac{1}{2^i} \frac{\|e_1f - e_1g\|_i}{1 + \|e_1f - e_1g\|_i} &\leq \sum_{i=1}^{N-1} \frac{1}{2^i} \frac{\|e_1f - e_1g\|_N}{1 + \|e_1f - e_1g\|_N} \\
&= \frac{\|e_1f - e_1g\|_N}{1 + \|e_1f - e_1g\|_N} \sum_{i=1}^{N-1} \frac{1}{2^i} \\
&< \frac{\|e_1f - e_1g\|_N}{1 + \|e_1f - e_1g\|_N}.
\end{aligned} \tag{4.15}$$

When inequality (4.12) holds, combining inequalities (4.11), (4.14), and (4.15) yields $d(h, f) < \varepsilon$. Hence, for any arbitrary $f \in \Gamma_0(\mathbb{R}^n)$ and $0 < \varepsilon < 1$, there exists a strongly convex function $h \in \Gamma_0(\mathbb{R}^n)$ such that $d(h, f) < \varepsilon$. That is, the set of strongly convex functions is dense in $(\Gamma_0(\mathbb{R}^n), d)$. Because $(\Gamma_0(\mathbb{R}^n), d)$ and $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$ are isometric by Corollary 3.8, it suffices to apply Lemma 2.23. The proof is complete. $\square$

Remark 4.15. *A shorter proof can be provided by approximating $f \in \Gamma_0(\mathbb{R}^n)$ with $f + \varepsilon\| \cdot \|^2$. One can use either [7, Theorem 7.4.5], or the fact that $f + \varepsilon\| \cdot \|^2$ converges uniformly to f on bounded subsets of* $\operatorname{dom} f$. *When $\varepsilon \downarrow 0$, clearly $f + \varepsilon\| \cdot \|^2$ converges epigraphically to f, which gives the Attouch-Wets convergence because we are in a finite dimension setting. To quantify the epigraphical convergence in terms of the distance d based on Moreau envelopes (Definition 3.3), one has*

$$e_1(f + \varepsilon\| \cdot \|^2) = (2\varepsilon + 1)e_1\left(\frac{f}{2\varepsilon + 1}\right)\left(\frac{\cdot}{2\varepsilon + 1}\right) + \frac{\varepsilon}{2\varepsilon + 1}\| \cdot \|^2,$$

which indeed converges to e_1f uniformly on bounded sets of $\mathbb{R}^n$.

Theorem 4.16. *The set of strongly convex functions is meager in $(e_1(\Gamma_0(\mathbb{R}^n)), \tilde{d})$ where $\tilde{d}$ is given by* (3.1). *Equivalently, in $(\Gamma_0(\mathbb{R}^n), d)$ the set of strongly convex function is meager.*

Proof. Denote the set of strongly convex functions in $e_1(\Gamma_0(\mathbb{R}^n))$ by S. Define

$$F_m := \left\{ g \in e_1(\Gamma_0(\mathbb{R}^n)) : g - \frac{1}{2m}\|\cdot\|^2 \text{ is convex on } \mathbb{R}^n \right\}.$$

We show that

a) $S = \bigcup_{m\in\mathbb{N}} F_m$,

b) for each $m \in \mathbb{N}$, the set F_m is closed in $e_1(\Gamma_0(\mathbb{R}^n))$, and

c) for each $m \in \mathbb{N}$, the set F_m has empty interior.

Then S will have been shown to be a countable union of closed, nowhere dense sets, hence first category.

a) $(\Rightarrow)$ Let $f \in S$. Then there exists $\sigma > 0$ such that $f - \frac{\sigma}{2}\|\cdot\|^2$ is convex. Note that this means $f - \frac{\tilde{\sigma}}{2}\|\cdot\|^2$ is convex for all $\tilde{\sigma} \in (0, \sigma)$. Since $\sigma > 0$, there exists $m \in \mathbb{N}$ such that $0 < \frac{1}{m} < \sigma$. Hence, $f - \frac{1}{2m}\|\cdot\|^2$ is convex, and $f \in F_m$. Therefore, $S \subseteq \bigcup_{m\in\mathbb{N}} F_m$.

$(\Leftarrow)$ Let $f \in F_m$ for some $m \in \mathbb{N}$. Then $f - \frac{1}{2m}\|\cdot\|^2$ is convex. Thus, with $\sigma = \frac{1}{m}$, we have that there exists $\sigma > 0$ such that $f - \frac{\sigma}{2}\|\cdot\|^2$ is convex, which is the definition of strong convexity of f. Therefore, $F_m \subseteq S$, and since this is true for every $m \in \mathbb{N}$, we have $\bigcup_{m\in\mathbb{N}} F_m \subseteq S$.

b) Let $g \notin F_m$. Then $g - \frac{1}{2m}\|\cdot\|^2$ is not convex. Equivalently, there exist $\lambda \in (0,1)$ and $x, y \in \mathbb{R}^n$ such that

$$\frac{g(\lambda x + (1-\lambda)y) - \lambda g(x) - (1-\lambda)g(y)}{\lambda(1-\lambda)} > -\frac{\|x-y\|^2}{2m}. \tag{4.16}$$

Let $N > \max\{\|x\|, \|y\|\}$. Choose $\varepsilon > 0$ such that when $\tilde{d}(f, g) < \varepsilon$ for $f \in e_1(\Gamma_0(\mathbb{R}^n))$, we have $\|f - g\|_N < \tilde{\varepsilon}$ for some $\tilde{\varepsilon} > 0$. In particular,

$$\begin{aligned}\frac{f(\lambda x + (1-\lambda)y) - \lambda f(x) - (1-\lambda)f(y)}{\lambda(1-\lambda)} =& \frac{g(\lambda x + (1-\lambda)y) - \lambda g(x) - (1-\lambda)g(y)}{\lambda(1-\lambda)} \\ &+ \frac{(f-g)(\lambda x + (1-\lambda)y) - \lambda(f-g)(x) - (1-\lambda)(f-g)(y)}{\lambda(1-\lambda)} \\ >& \frac{g(\lambda x + (1-\lambda)y) - \lambda g(x) - (1-\lambda)g(y)}{\lambda(1-\lambda)} - \frac{4\tilde{\varepsilon}}{\lambda(1-\lambda)}.\end{aligned}$$

Hence, when $\tilde{\varepsilon}$ is sufficiently small, which can be achieved by making ε sufficiently small, we have

$$\frac{f(\lambda x + (1-\lambda)y) - \lambda f(x) - (1-\lambda)f(y)}{\lambda(1-\lambda)} > -\frac{\|x-y\|^2}{2m}.$$

This gives us, by equation (4.16), that $f - \frac{1}{2m}\|\cdot\|^2$ is not convex. Thus, $f \notin F_m$, so $e_1(\Gamma_0(\mathbb{R}^n)) \setminus F_m$ is open, and therefore F_m is closed.

c) That $\operatorname{int} F_m = \emptyset$ is equivalent to saying that $e_1(\Gamma_0(\mathbb{R}^n)) \setminus F_m$ is dense. Thus, it suffices to show that for every $\varepsilon > 0$ and every $g \in e_1(\Gamma_0(\mathbb{R}^n))$, the open ball $\mathbb{B}_\varepsilon(g)$ contains an element of $e_1(\Gamma_0(\mathbb{R}^n)) \setminus F_m$.
If $g \in e_1(\Gamma_0(\mathbb{R}^n)) \setminus F_m$, then there is nothing to prove. Assume that $g \in F_m$. Then g is $\frac{1}{2m}$-strongly convex, and has a strong minimizer $\bar{x}$ by Lemma 2.26. As $g \in e_1(\Gamma_0(\mathbb{R}^n))$, $g = e_1 f$ for some $f \in \Gamma_0(\mathbb{R}^n)$. We consider two cases.

Case 1: Suppose that for every $\frac{1}{k} > 0$, there exists $x_k \neq \bar{x}$ such that $f(x_k) < f(\bar{x}) + \frac{1}{k}$. Define $h_k := \max\left\{f, f(\bar{x}) + \frac{1}{k}\right\}$. Then

$$\min h_k = f(\bar{x}) + \frac{1}{k},\ f \leq h_k < f + \frac{1}{k},$$

so that $e_1 f \leq e_1 h_k \leq e_1 f + \frac{1}{k}$. We have $g_k := e_1 h_k \in e_1(\Gamma_0(\mathbb{R}^n))$, and $\|g_k - g\|_i < \frac{1}{k}$ for all $i \in \mathbb{N}$. Choosing k sufficiently large guarantees that $\tilde{d}(g_k, g) < \varepsilon$. We see that g_k does not have a strong minimizer by noting that for every k, $f(\bar{x}) < f(\bar{x}) + \frac{1}{k}$, $f(x_k) < f(\bar{x}) + \frac{1}{k}$, and $h_k(\bar{x}) = h_k(x_k) = f(\bar{x}) + \frac{1}{k}$. Thus, h_k does not have a strong minimizer, which implies that $g_k = e_1 h_k$ does not either, by Proposition 4.1. Therefore, $g_k \notin F_m$.

Case 2: If Case 1 is not true, then there exists k such that $f(x) \geq f(\bar{x}) + \frac{1}{k}$ for every $x \neq \bar{x}$. Then we claim that $f(x) = \infty$ for all $x \neq \bar{x}$. Suppose for the purpose of contradiction that there exists $x \neq \bar{x}$ such that $f(x) < \infty$. As $f \in \Gamma_0(\mathbb{R}^n)$, the function $\phi : [0,1] \to \mathbb{R}$ defined by $\phi(t) := f(tx + (1-t)\bar{x})$ is continuous by [32, Proposition 2.1.6]. This contradicts the assumption, therefore,

$$f(x) = \iota_{\{\bar{x}\}}(x) + f(\bar{x}).$$

Consequently,

$$g(x) = e_1 f(x) = f(\bar{x}) + \frac{1}{2}\|x - \bar{x}\|^2.$$

Now for every $j \in \mathbb{N}$, define $f_j : \mathbb{R}^n \to \overline{\mathbb{R}}$,

$$f_j(x) := \begin{cases} f(\bar{x}), & \|x - \bar{x}\| \leq \frac{1}{j}, \\ \infty, & \text{otherwise.} \end{cases}$$

We have $f_j \in \Gamma_0(\mathbb{R}^n)$, and

$$g_j(x) := e_1 f_j(x) = \begin{cases} f(\bar{x}), & \|x - \bar{x}\| \leq \frac{1}{j}, \\ f(\bar{x}) + \frac{1}{2}\left(\|x - \bar{x}\| - \frac{1}{j}\right)^2, & \|x - \bar{x}\| > \frac{1}{j}. \end{cases}$$

Then $\{g_j(x)\}_{j \in \mathbb{N}}$ converges pointwise to $e_1 f = g$, by [27, Theorem 7.37]. Thus, for sufficiently large j, $\tilde{d}(g_j, g) < \varepsilon$. Since g_j is constant on $\mathbb{B}_{\frac{1}{j}}(\bar{x})$, g_j is not strongly convex, so $g_j \notin F_m$.

Properties a), b) and c) all together show that the set of strongly convex function is meager in $(e_1(\Gamma_0(\mathbb{R}^n), \tilde{d})$. Note that $(e_1(\Gamma_0(\mathbb{R}^n), \tilde{d})$ and $(\Gamma_0(\mathbb{R}^n), d)$ are isometric by Corollary 3.8. The proof is complete by using Lemma 2.23. $\square$

Remark 4.17. *(1). Theorem 4.16 shows that the set S of strongly convex functions is first category, but what about the larger set U of uniformly convex functions to which S belongs? (See [5, 31, 32] for information on uniform convexity.) It is clear that U is bigger than S, for example [5, Definition 10.5] shows that $f \in S \Rightarrow f \in U$, and [5, Exercise 10.7] states that $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|^4$ is uniformly convex but not strongly convex. So is U generic or meager? This is a question for which the authors have no answer as of yet; it remains an open question. We refer the reader in particular to [32, Proposition 3.5.8, Theorem 3.5.10] for properties and characterizations of uniformly convex functions.*

(2). Strongly convex functions are important in optimization [5]. In [25], Rockafellar showed that the proximal point method associated to every strongly convex function converges at least at a linear rate to the optimal solution.

Corollary 4.18. *In $(\Gamma_0(\mathbb{R}^n), d))$ the set*

$$D := \{f : f \text{ is differentiable and } \nabla f \text{ is } c\text{-Lipschitz for some } c > 0\}$$

is first category.

Proof. We know that $(\Gamma_0(\mathbb{R}^n), d) \to ((\Gamma_0(\mathbb{R}^n))^*, \hat{d})$ is an isometry, and that f is strongly convex for some $c > 0$ if and only if f^* is differentiable on $\mathbb{R}^n$ with ∇f^* being $\frac{1}{c}$-Lipschitz [32, Corollary 3.5.11(i)⇔(vi)]. Since the set of strongly convex functions is first category in $((\Gamma_0(\mathbb{R}^n))^*, \hat{d})$, the set of differentiable convex functions with ∇f being c-Lipschitz for some $c > 0$ is a first category set in $(\Gamma_0(\mathbb{R}^n), d)$. $\square$

4.3 The set of convex functions with strong minimizers is second category

We present properties of the sets U_m and E_m, and show that the set of convex functions that attain a strong minimum is a generic set in $(\Gamma_0(\mathbb{R}^n), d)$.

Lemma 4.19. *The sets U_m and E_m are dense in $(\Gamma_0(\mathbb{R}^n), d)$.*

Proof. This is immediate by combining Theorems 4.12 and 4.14. $\square$

To continue, we need the following result, which holds in $\Gamma_0(X)$ where X is any Banach space.

Lemma 4.20. *Let $f \in \Gamma_0(\mathbb{R}^n)$, $m \in \mathbb{N}$, and fix $z \in \operatorname{dom} f$. Then*

$$\inf_{\|x-z\| \geq \frac{1}{m}} f(x) - f(z) > 0 \text{ if and only if } \inf_{m \geq \|x-z\| \geq \frac{1}{m}} f(x) - f(z) > 0.$$

Proof. $(\Rightarrow)$ Suppose that for z fixed, $\inf\limits_{\|x-z\|\geq\frac{1}{m}} f(x)-f(z)>0$. Since

$$\inf_{m\geq\|x-z\|\geq\frac{1}{m}} f(x)-f(z)\geq \inf_{\|x-z\|\geq\frac{1}{m}} f(x)-f(z),$$

we have $\inf\limits_{m\geq\|x-z\|\geq\frac{1}{m}} f(x)-f(z)>0$.

$(\Leftarrow)$ Let $\inf\limits_{m\geq\|x-z\|\geq\frac{1}{m}} f(x)-f(z)>0$, and suppose that

$$\inf_{\|x-z\|\geq\frac{1}{m}} f(x)-f(z)\leq 0.$$

Then for each $\frac{1}{k}$ with $k\in\mathbb{N}$, there exists y_k with $\|y_k-z\|\geq\frac{1}{m}$ such that $f(y_k)\leq f(z)+\frac{1}{k}$. Take $z_k\in[y_k,z]\cap\left\{x\in\mathbb{R}^n:\ m\geq\|x-z\|\geq\frac{1}{m}\right\}\neq\emptyset$. Then

$$z_k=\lambda_k y_k+(1-\lambda_k)z$$

for some $\lambda_k\in[0,1]$. By the convexity of f, we have

$$\begin{aligned}
f(z_k)=f(\lambda_k y_k+(1-\lambda_k)z)&\leq\lambda_k f(y_k)+(1-\lambda_k)f(z)\\
&\leq\lambda_k f(z)+(1-\lambda_k)f(z)+\frac{\lambda_k}{k}\\
&=f(z)+\frac{\lambda_k}{k}\leq f(z)+\frac{1}{k}.
\end{aligned}$$

Now $\inf\limits_{m\geq\|x-z\|\geq\frac{1}{m}} f(x)\leq f(z_k)\leq f(z)+\frac{1}{k}$, so when $k\to\infty$ we obtain

$$\inf_{m\geq\|x-z\|\geq\frac{1}{m}} f(x)-f(z)\leq 0.$$

This contradicts the fact that $\inf\limits_{m\geq\|x-z\|\geq\frac{1}{m}} f(x)-f(z)>0$. Therefore, $\inf\limits_{\|x-z\|\geq\frac{1}{m}} f(x)-f(z)>0$. $\square$

Lemma 4.21. *The set E_m is an open set in $(\Gamma_0(\mathbb{R}^n),d)$.*

Proof. Fix $m\in\mathbb{N}$, and let $f\in E_m$. Then there exists $z\in\mathbb{R}^n$ such that $\inf\limits_{\|x-z\|\geq\frac{1}{m}} e_1f(x)-e_1f(z)>0$. Hence, by Lemma 4.20,

$$\inf_{m\geq\|x-z\|\geq\frac{1}{m}} e_1f(x)-e_1f(z)>0.$$

Choose j large enough that $\mathbb{B}_m[z]\subseteq\mathbb{B}_j(0)$. Let $g\in\Gamma_0(\mathbb{R}^n)$ be such that $d(f,g)<\varepsilon$, where

$$0<\varepsilon<\frac{\inf\limits_{m\geq\|x-z\|\geq\frac{1}{m}} e_1f(x)-e_1f(z)}{2^j\left(2+\inf\limits_{m\geq\|x-z\|\geq\frac{1}{m}} e_1f(x)-e_1f(z)\right)}<\frac{1}{2^j}. \tag{4.17}$$

The reason for this bound on ε will become apparent at the end of the proof. Then

$$\sum_{i=1}^{\infty}\frac{1}{2^i}\frac{\|e_1 f - e_1 g\|_i}{1+\|e_1 f - e_1 g\|_i} < \varepsilon.$$

In particular for our choice of j, we have that $2^j\varepsilon < 1$ by inequality (4.17), and that

$$\begin{aligned}
\frac{1}{2^j}\frac{\|e_1 f - e_1 g\|_j}{1+\|e_1 f - e_1 g\|_j} &< \varepsilon,\\
\|e_1 f - e_1 g\|_j &< 2^j\varepsilon(1+\|e_1 f - e_1 g\|_j),\\
\sup_{\|x\|\leq j}|e_1 f(x) - e_1 g(x)|(1-2^j\varepsilon) &< 2^j\varepsilon,\\
\sup_{\|x\|\leq j}|e_1 f(x) - e_1 g(x)| &< \frac{2^j\varepsilon}{1-2^j\varepsilon}.
\end{aligned}$$

Define $\alpha := \frac{2^j\varepsilon}{1-2^j\varepsilon}$. Then $\sup_{\|x\|\leq j}|e_1 f(x) - e_1 g(x)| < \alpha$. Hence,

$$|e_1 f(x) - e_1 g(x)| < \alpha \text{ for all } x \text{ with } \|x\| \leq j.$$

In other words,

$$e_1 f(x) - \alpha < e_1 g(x) < e_1 f(x) + \alpha \text{ for all } x \text{ with } \|x\| \leq j.$$

Since $\mathbb{B}_m[z] \subseteq \mathbb{B}_j(0)$, we can take the infimum over $m \geq \|x-z\| \geq \frac{1}{m}$ to obtain

$$\inf_{m\geq\|x-z\|\leq\frac{1}{m}} e_1 f(x) - \alpha \leq \inf_{m\geq\|x-z\|\geq\frac{1}{m}} e_1 g(x) \leq \inf_{m\geq\|x-z\|\geq\frac{1}{m}} e_1 f(x) + \alpha. \tag{4.18}$$

Using inequality (4.18) together with the fact that $|e_1 g(z) - e_1 f(z)| < \alpha$ yields

$$\begin{aligned}
\inf_{m\geq\|x-z\|\geq\frac{1}{m}} e_1 g(x) - e_1 g(z) &\geq \left(\inf_{m\geq\|x-z\|\geq\frac{1}{m}} e_1 f(x) - \alpha\right) - (e_1 f(z) + \alpha)\\
&= \inf_{m\geq\|x-z\|\geq\frac{1}{m}} e_1 f(x) - e_1 f(z) - 2\alpha.
\end{aligned}$$

Hence, if

$$\alpha < \frac{\inf_{m\geq\|x-z\|\geq\frac{1}{m}} e_1 f(x) - e_1 f(z)}{2}, \tag{4.19}$$

we have

$$\inf_{m\geq\|x-z\|\geq\frac{1}{m}} e_1 g(x) - e_1 g(z) > 0. \tag{4.20}$$

Recalling that $\alpha = \frac{2^j\varepsilon}{1-2^j\varepsilon}$, we solve equation (4.19) for ε to obtain

$$\varepsilon < \frac{\inf\limits_{m\geq\|x-z\|\geq\frac{1}{m}} e_1f(x) - e_1f(z)}{2^j\left(2 + \inf\limits_{m\geq\|x-z\|\geq\frac{1}{m}} e_1f(x) - e_1f(z)\right)}.$$

Thus, inequality (4.20) is true whenever $d(f,g) < \varepsilon$ for any ε that respects inequality (4.17). Applying Lemma 4.20 to inequality (4.20), we conclude that

$$\inf_{\|x-z\|\geq\frac{1}{m}} e_1g(x) - e_1g(z) > 0.$$

Hence, if $g \in \Gamma_0(\mathbb{R}^n)$ is such that $d(f,g) < \varepsilon$, then $g \in E_m$. Therefore, E_m is open. $\square$

We are now ready to present the main results of the paper.

Theorem 4.22. *In $X := (\Gamma_0(\mathbb{R}^n), d)$, the set $S := \{f \in \Gamma_0(\mathbb{R}^n) : f$ attains a strong minimum$\}$ is generic.*

Proof. By Lemmas 4.19 and 4.21, we have that E_m is open and dense in X. Hence, $G := \bigcap\limits_{m\in\mathbb{N}} E_m$ is a countable intersection of open, dense sets in X, and as such G is generic in X. Let $f \in G$. By Corollary 4.11, f attains a strong minimum on $\mathbb{R}^n$. Thus, every element of G attains a strong minimum on $\mathbb{R}^n$. Since G is generic in X and $G \subseteq S$, we conclude that S is generic in X. $\square$

Remark 4.23. *This result is stated as Exercise 7.5.10 in [7, p. 269]. However, the approach taken there uses the Attouch-Wets topology defined by uniform convergence on bounded subsets of the distance function, associated with epigraphs of convex functions.*

Theorem 4.24. *In $X := (\Gamma_0(\mathbb{R}^n), d)$, the set $S := \{f \in \Gamma_0(\mathbb{R}^n) : f$ is coercive$\}$ is generic.*

Proof. Define the set $\Gamma_1(\mathbb{R}^n) := \Gamma_0(\mathbb{R}^n) + x^*$, in the sense that for any function $f \in \Gamma_0(\mathbb{R}^n)$, the function $f + \langle x^*, \cdot\rangle \in \Gamma_1(\mathbb{R}^n)$. Since any such $f + \langle x^*, \cdot\rangle$ is proper, lsc, and convex, we have $\Gamma_1(\mathbb{R}^n) \subseteq \Gamma_0(\mathbb{R}^n)$. Now, since for any $f \in \Gamma_0(\mathbb{R}^n)$ we have that $f - x^* \in \Gamma_0(\mathbb{R}^n)$, this gives us that $f \in \Gamma_0(\mathbb{R}^n) + x^* = \Gamma_1(\mathbb{R}^n)$. Therefore, $\Gamma_1(\mathbb{R}^n) = \Gamma_0(\mathbb{R}^n)$. By Theorem 4.22, there exists a generic set $G \subseteq \Gamma_0(\mathbb{R}^n)$ such that for every $f \in G$, f attains a strong minimum at some point x, and hence $0 \in \partial f(x)$. Then, given any x^* fixed, there exists a generic set G_{x^*} that contains a dense G_δ set, such that $0 \in \partial(f + x^*)(x)$. Thus, for each $f \in G_{x^*}$ there exists $x \in \mathbb{R}^n$ such that $-x^* \in \partial f(x)$. By Fact 2.4, it is possible to construct the set $D := \{-x_i^*\}_{i=1}^\infty$ such that $\overline{D} = \mathbb{R}^n$. Then each set $G_{x_i^*}$, $i \in \mathbb{N}$, contains a dense G_δ set. Therefore, the set $G := \bigcap\limits_{i=1}^\infty G_{x_i^*}$ contains a dense G_δ set. Let $f \in G$. Then for each $i \in \mathbb{N}$, $-x_i^* \in \partial f(x)$ for some $x \in \mathbb{R}^n$. That is, $-x_i^* \in \operatorname{ran}\partial f$. So $D := \bigcup\limits_{i=1}^\infty\{-x_i^*\} \subseteq \operatorname{ran}\partial f$, and $\overline{D} \subseteq \overline{\operatorname{ran}\partial f}$. Since $\overline{D} = \mathbb{R}^n$, we have $\mathbb{R}^n = \overline{\operatorname{ran}\partial f}$. By Facts 2.5 and 2.6, $\operatorname{ran}\partial f$ is almost convex; there exists a convex set C such that $C \subseteq \operatorname{ran} f \subseteq \overline{C}$.

Then $\overline{C} = \mathbb{R}^n$. As C is convex, by [26, Theorem 6.3] we have the relative interior $\operatorname{ri}\overline{C} = \operatorname{ri} C$, so $\operatorname{ri} C = \mathbb{R}^n$. Thus, $\mathbb{R}^n = \operatorname{ri} C \subseteq C$, which gives us that $C = \mathbb{R}^n$. Therefore, $\operatorname{ran}\partial f = \mathbb{R}^n$. By Fact 2.8, $\operatorname{ran}\partial f \subseteq \operatorname{dom}(f^*)$. Hence, $\operatorname{dom} f^* = \mathbb{R}^n$. By Fact 2.22, we have that $\lim_{\|x\|\to\infty} \frac{f(x)}{\|x\|} = \infty$. Therefore, f is coercive for all $f \in G$. Since G is generic in X and $G \subseteq S$, we conclude that S is generic in X. □

Theorem 4.25. *In* $(\Gamma_0(\mathbb{R}^n), d)$*, the set* $S := \{f \in \Gamma_0(\mathbb{R}^n) : \operatorname{dom} f = \mathbb{R}^n\}$ *is generic.*

Proof. Note that $(\Gamma_0(\mathbb{R}^n))^* = \Gamma_0(\mathbb{R}^n)$. In $((\Gamma_0(\mathbb{R}^n))^*, d)$, by Theorem 4.24, the set

$$\{f^* \in (\Gamma_0(\mathbb{R}^n))^* : f^* \text{ is coercive}\}$$

is generic. Since f^* is coervcive if and only if f has $\operatorname{dom} f = \mathbb{R}^n$ by Fact 2.22, the proof is done. □

Combining Theorems 4.22, 4.24 and 4.25, we obtain

Corollary 4.26. *In* $(\Gamma_0(\mathbb{R}^n), d)$*, the set*

$$S := \{f \in \Gamma_0(\mathbb{R}^n) : \operatorname{dom} f = \mathbb{R}^n, \operatorname{dom} f^* = \mathbb{R}^n, f \textit{ has a strong minimizer}\}$$

is generic.

5 Conclusion

Endowed with the Attouch-Wets metric, based on the Moreau envelope, the set of proper lower semicontinuous convex functions on finite-dimensional space forms a complete metric space. In this complete metric space, the topology is epi-convergence topology. We have proved several Baire category results. In particular, we have shown that in $(\Gamma_0(\mathbb{R}^n), d)$ the set of strongly convex functions is category one, the set of functions that attain a strong minimum is category two, and the set of coercive functions is category two. Several other results about strongly convex functions and functions with strong minima are included. In future work that has already commenced, we will continue to develop the theory of Moreau envelopes, providing characterizations and illustrative examples of how to calculate them, and extend results in this paper to convex functions defined on Hilbert spaces and to prox-bounded functions on $\mathbb{R}^n$. A natural question to ask here, as a referee did, are these genericity results sharp? According to [32, Proposition 3.5.8, Theorem 3.5.10], a uniformly convex function is coercive, has a strong minimizer, and its Fenchel conjugate f^* is differentiable with ∇f^* being uniformly continuous. Can more conclusions be drawn, for instance about the set of uniformly convex functions of which the set of strongly convex functions is a subset? The authors do not have the answer, and will consider the question in future research.

Acknowledgements

The authors wish to thank the two anonymous referees for their valuable suggestions for improvement of this paper, in particular, Remarks 2.16, 3.6, 4.15. The results of their comments were a better overall presentation and a broadened connection to existing literature on the topic.

References

[1] H. Attouch. *Variational convergence for functions and operators*. Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[2] H. Attouch, R. Lucchetti, and R. J. Wets. The topology of the ρ-Hausdorff distance. *Ann. Mat. Pura Appl. (4)*, 160:303–320, 1991.

[3] H. Attouch and R. J. Wets. Isometries for the Legendre-Fenchel transform. *Trans. Amer. Math. Soc.*, 296(1):33–60, 1986.

[4] J.-B. Baillon and G. Haddad. Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones. *Israel J. Math.*, 26, 1977.

[5] H. H. Bauschke and P. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2011.

[6] H. H. Bauschke, R. Goebel, Y. Lucet, and X. Wang. The proximal average: basic theory. *SIAM J. Optim.*, 19(2):766–785, 2008.

[7] G. Beer. *Topologies on closed and closed convex sets*. Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht, 1993.

[8] G. Beer and R. Lucchetti. Convex optimization and the epi-distance topology. *Trans. Amer. Math. Soc.*, 327(2):795–813, 1991.

[9] G. Beer and R. Lucchetti. The epi-distance topology: Continuity and stability results with applications to convex optimization problems. *Math. Oper. Res.*, 17(3):715–726, 1992.

[10] J. M. Borwein and A. S. Lewis. *Convex analysis and nonlinear optimization*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 3. Springer, New York, second edition, 2006.

[11] J. M. Borwein and J. D. Vanderwerff. *Convex functions: constructions, characterizations and counterexamples*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2010.

[12] J. M. Borwein and Q. Zhu. *Techniques of variational analysis*. CMS Books in Mathematics. Springer-Verlag, New York, 2005.

[13] M. M. Čoban, P. S. Kenderov, and J. P. Revalski. Generic well-posedness of optimization problems in topological spaces. *Mathematika*, 36(2):301–324 (1990), 1989.

[14] R. Deville, G. Godefroy, and V. Zizler. A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions. *J. Funct. Anal.*, 111(1):197–212, 1993.

[15] R. Deville and J. P. Revalski. Porosity of ill-posed problems. *Proc. Amer. Math. Soc.*, 128(4):1117–1124, 2000.

[16] A. L. Dontchev and T. Zolezzi. *Well-posed optimization problems*. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1993.

[17] P. S. Kenderov and J. P. Revalski. The Banach-Mazur game and generic existence of solutions to optimization problems. *Proc. Amer. Math. Soc.*, 118(3):911–917, 1993.

[18] S. Lacoste-Julien, M. Schmidt, and F. Bach. A simpler approach to obtaining an O(1/t) convergence rate for the projected stochastic subgradient method. *arXiv:1212.2002*, pages 1–8, 2012.

[19] R. Lucchetti. *Convexity and well-posed problems*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 22. Springer, New York, 2006.

[20] J.-J. Moreau. Propriétés des applications "prox". *C. R. Acad. Sci. Paris*, 256:1069–1071, 1963.

[21] J.-J. Moreau. Proximitéet dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

[22] C. Planiden and X. Wang. Most convex functions have unique minimizers. *J. Convex Anal.*, to appear, 2016.

[23] S. Reich and A. J. Zaslavski. *Genericity in nonlinear analysis*. Developments in Mathematics. Springer, New York, 2014.

[24] J. P. Revalski and N. V. Zhivkov. Well-posed constrained optimization problems in metric spaces. *J. Optim. Theory Appl.*, 76(1):145–163, 1993.

[25] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optimization*, 14(5):877–898, 1976.

[26] R. T. Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997.

[27] R. T. Rockafellar and R. J. Wets. *Variational analysis*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1998.

[28] J. E. Spingarn and R. T. Rockafellar. The generic nature of optimality conditions in nonlinear programming. *Math. Oper. Res.*, 4(4):425–430, 1979.

[29] X. Wang. Most maximally monotone operators have a unique zero and a super-regular resolvent. *Nonlinear Anal.*, 87:69–82, 2013.

[30] K. Yosida. *Functional analysis*. Die Grundlehren der Mathematischen Wissenschaften, Band 123. Academic Press, Inc., New York; Springer-Verlag, Berlin, 1965.

[31] C. Zălinescu. On uniformly convex functions. *J. Math. Anal. Appl.*, 95(2):344 – 374, 1983.

[32] C. Zălinescu. *Convex analysis in general vector spaces*. World Scientific Publishing Co., Inc., River Edge, NJ, 2002.